## Supplemental Data

HOXB1 Founder Mutation in Humans

Recapitulates the Phenotype of *Hoxb1*$^{-/-}$ Mice

Bryn D. Webb, Sherin Shaaban, Harald Gaspar, Luis F. Cunha, Christian R. Schubert, Ke Hao, Caroline D. Robson, Wai-Man Chan, Caroline Andrews, Sarah MacKinnon, Darren T. Oystreck, David G. Hunter, Anthony J. Iacovelli, Xiaoqian Ye, Anne Camminady, Elizabeth C. Engle, Ethylin Wang Jabs

Figure S1. 3D morphometric analysis and clinical summary
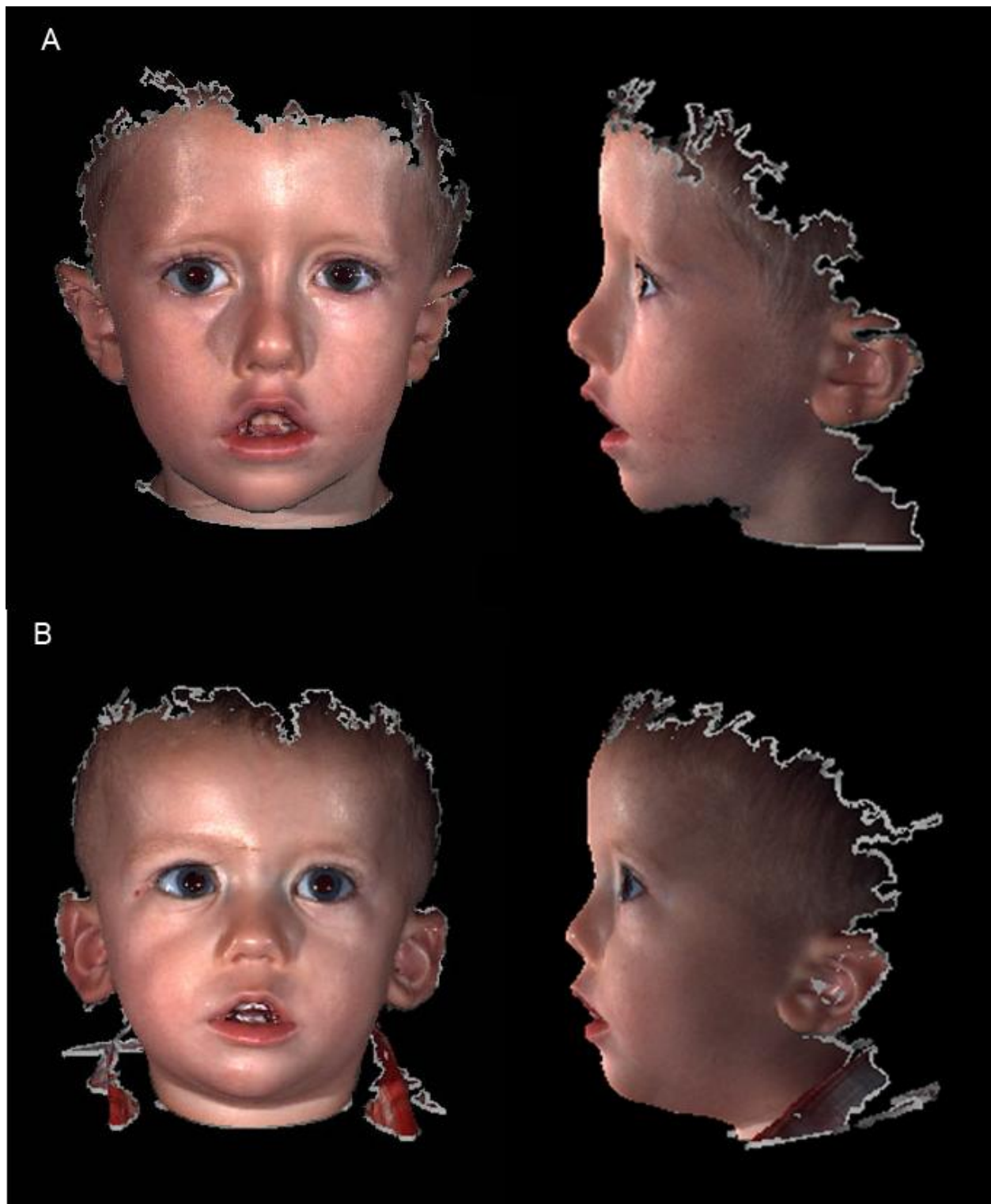
(A-B) 3-Dimensional photographs were obtained for the affected children in Family A using the 3DMD system (http://3dmd.com). A 3DMD image is acquired using multiple cameras and takes approximately 1 to 2 seconds to obtain. Measurements were obtained using a freehand point selection option.

The older affected male in Family A (Patient VI-1 in Figure 2A) was examined at 5 yrs and 7.4 months (A). His head is normocephalic at the 50th to 75th percentiles (clinical exam measurement, head circumference is 52.5 cm).[1] He has normal distribution of hair. Eyebrows are unremarkable, and there is no obvious hypo- or hypertelorism. Inner canthal distance measures 2.7 cm (25th -50th percentile) and outer canthal distance measures 8.1 cm (50th -75th percentile).[2] Right palpebral fissure measures 2.6 cm (-1SD to mean) and left palpebral fissure measures 2.7 cm (mean to +1SD).[1] Ears are well-formed, but are low-set and posteriorly rotated. Right ear measures 5.2 cm (-1SD to mean) and left ear measures 5.1 cm (1SD below mean).[1] He has midface retrusion. His nasal tip is slightly upturned and nasal bridge is prominent. Nasal length is 4.2 cm (+1SD above mean);[3] nasal protrusion is 1.7 cm (+1 to +2SD above mean).[3] Nasal width measures 2.4 cm (>>2SD below mean).[1] The patient has a short, smooth philtrum. Philtrum length is 0.9 cm (>2SD below mean).[3] Mouth is open. Intercommisural distance is 3.6 cm (-2 to -1SD below mean).[1] Chin appears mildly micrognathic, but not distinctly abnormal. Facial index is 85.7 (mesoprosopic).[4]

The younger affected male in Family A (Patient VI-3 in Figure 2A) was examined at 1 yr and 5 months (B). His head is normocephalic at the 50th to 75th percentiles (clinical exam measurement, head circumference is 49.3 cm).[1] He has normal distribution of hair. He has a broad forehead. Eyebrows are unremarkable, and eye examination is notable for epicanthal folds. There is no obvious hypo- or hypertelorism. Inner canthal distance measures 2.5 cm (25th -50th percentile) and outer canthal distance measures 7.5 cm (50th to 75th percentile).[2] Right palpebral fissure measures 2.4 cm (+1SD above mean) and left palpebral fissure measures 2.5

cm (+1 to +2SD above mean).[1] Ears are well-formed, but are low-set and posteriorly rotated. Right ear measures 5.1 cm (+1 to +2SD above mean) and left ear measures 5.2 cm (+1 to +2SD above mean).[1] He has midface retrusion. His nasal tip is slightly upturned. Nasal length is 3.2 cm (-2 to -1SD below mean);[1] nasal protrusion is 1.2 cm (mean).[3] Nasal width measures 2.5 cm (-2 to -1SD below mean).[1] The patient has a smooth philtrum. Philtrum length is 1.1 cm (25th percentile).[1] Mouth is open. Intercommisural distance is 3.4 cm (mean to +1SD).[1] Chin is not abnormal. Facial index is 77.9 (euryprosopic).[4]

Both boys were born at full term following uneventful pregnancies with no known *in utero* exposures. Individual VI-3 had torticollis and mild gross motor delay.VI-1 has normal developmental and school performances, and has mild intermittent head tics. Both boys exam revealed normal anterior and posterior segments of the eye and intact pupillary reactions. Sense of smell is normal, as is facial sensation in V1-V3 distribution. They are unable to smile, purse their lips, or close their eyes against resistance, and because of lagophthalmos they require eye lubrication during sleep. They have minimal forehead wrinkling when crying. Taste discrimination, salivation, and lacrimation are intact, as is general sensation over the concha of the ear and skin behind the auricle. Soft and hard palates are intact, gag reflex is present, and swallowing is normal, but both children have some difficulties chewing secondary to facial weakness. Tongues are midline with normal strength and movement. Muscles other than facial are of full strength, deep tendon reflexes are brisk but without clonus, no pathological reflexes are present, and sensation, coordination, and gait are normal. Cardiac, respiratory, GI, GU, and skin exams are normal, and neither child has evidence of scoliosis, vertebral, or limb anomalies. VI-1 underwent an EEG for evaluation of tics which was reportedly normal. Echocardiography was normal in both children.

Figure S2. Family structure and genetic analysis of Family A
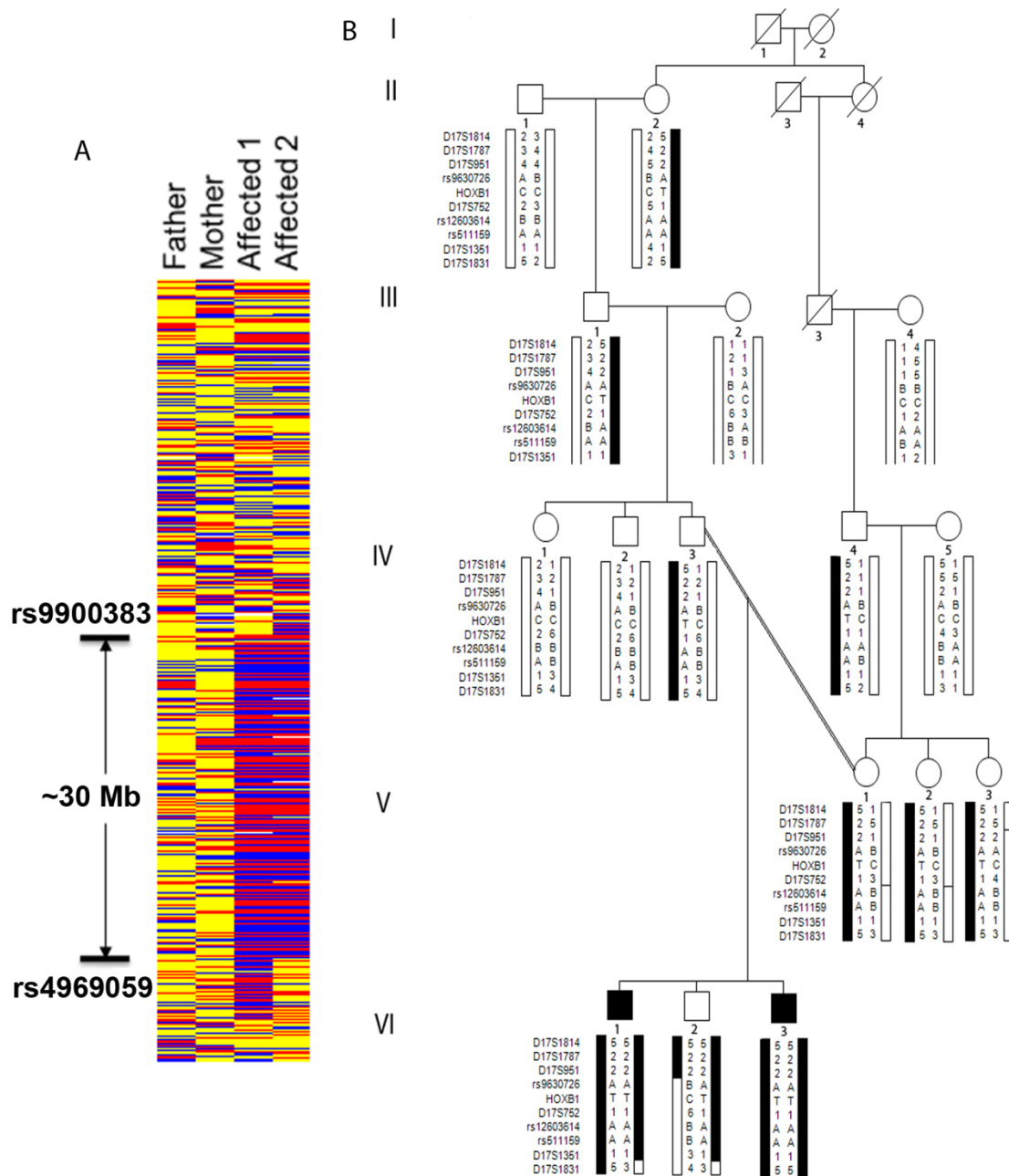
(A) Schematic of homozygosity mapping using dChip software.[5] Genotypes created using

Genome-Wide Human SNP Array 6.0 (Affymetrix Inc., www.affymetrix.com) from individuals IV-

3 (father), V-1 (mother), VI-1 (affected 1) and VI-2 (affected 2) as per pedigree in (B) are displayed. Genotypes are homozygous if red or blue, heterozygous if yellow and absent if white. The overlapping homozygous region on chromosome 17 between the two affected sibs was about 30Mb bordered by markers rs9900383 and rs4969059. For linkage and haplotype analysis, genomic DNA was extracted from peripheral blood samples using the Puregene kit (Qiagen, Valencia, CA) or from saliva using Oragene Discover DNA Self-Collection Kits (DNA Genotek). Genotyping was performed using Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix Inc., www.affymetrix.com), and SNP genotypes and copy number variants were called using Birdsuite v2[6] set of algorithms. We assumed reduction to homozygosity surrounding the causative gene mutation in the affected children,[7] and such regions were identified using dChip software.[5] (B) Schematic of Family A pedigree structure and mutation status for 16 family members who were enrolled during the course of the study. Squares denote males, circles denote females, and filled symbols denote affection. Double bars represent consanguinity. Linkage of the extended family was performed using PCR amplification of fluorescently labeled polymorphic microsatellite markers spanning the chromosome 17 region of homozygosity. The amplified products were analyzed on an ABI 3730 DNA sequencer (XL Genetic Analyzer) and alleles were assigned with GeneMapper v.3.5 (Applied Biosystems). For linkage analysis, we assumed full penetrance without phenocopies, equal recombination frequencies in males and females, and an autosomal recessive mode of inheritance with a gene frequency of 1:100,000.

Table S1. Exome sequence data.

	Variants	Non-pathogenic	Potentially pathogenic			
		Synonomous	Non-synonomous missense	Nonsense	Indels	Splice-sites
Total	19625	10174	8746	114	506	85
Not in dbSNP	878	206	330	15	300	27
And not in 1000 genome	615	140	242	12	199	22
And homozygous	53	7	19	1	21	5
And not in EVS	26	2	11	0	8	5
And in region of homozygosity	6	1	5*	0	0	0

*Variants were found in 5 genes: *HOXB1*, *CA4*, *USP32*, *C17orf57* and *WFIKKN2*

Exome enrichment of 3 µg genomic DNA from VI-1 was performed by the Broad Institute using the SureSelect Human All Exon Kit (Agilent, Santa Clara, CA, USA), and the enriched library was amplified and sequenced on Illumina HiSeq 2000 (Illumina, San Diego, CA, USA). Paired-end sequences were created at a read length of 100bp and then mapped and aligned to the human reference genome (NCBI build 27.2) using MAQ/BWA tools.[8-9] An average coverage of 90.4% at 20X was obtained for the 32Mb of captured exome sequence. Variant calling of single nucleotide polymorphisms (SNPs) and insertions/deletions was done with GATK/DINDEL tools respectively[10-11] and annotated using Annovar annotation Package.[12] Exome variants were filtered based on: (a) their absence in dbSNP132[13], 1000 Genomes Project[14], and a frequency <1% in the 5379 exomes entered into the Exome Variant Server, NHLBI Exome Sequencing Project (ESP), Seattle, WA (URL:http://evs.gs.washington.edu/EVS/); (b) homozygosity in the affected child; (c) location within region of homozygosity.

References:

1. Hall, J., Froster-Iskenius, U., and Allanson, J. (1989). Handbook of normal physical measurements (Oxford: Oxford University Press).

2. Saul, R., Geer, J., Seaver, L., Phelen, M., Sweet, K., and Mills, C. (1998). Growth References: Third Trimester to Adulthood (Greenwood, South Carolina: Greenwood Genetic Center).

3. Zankl, A., Eberle, L., Molinari, L., and Schinzel, A. (2002). Growth charts for nose length, nasal protrusion, and philtrum length from birth to 97 years. Am J Med Genet 111, 388-391.

4. Banister, M. (1995). Skeletal system. In Gray's Anatomy, P. Williams, M. Dyson, J. Dussak, and et al., eds. (London, Churchill Livingstone), pp 607-612.

5. Lin, M., Wei, L.J., Sellers, W.R., Lieberfarb, M., Wong, W.H., and Li, C. (2004). dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. Bioinformatics 20, 1233-1240.

6. Korn, J.M., Kuruvilla, F.G., McCarroll, S.A., Wysoker, A., Nemesh, J., Cawley, S., Hubbell, E., Veitch, J., Collins, P.J., Darvishi, K., et al. (2008). Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat Genet 40, 1253-1260.

7. Lander, E.S., and Botstein, D. (1987). Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science 236, 1567-1570.

8. Li, H., Ruan, J., and Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 18, 1851-1858.

9. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754-1760.

10. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43, 491-498.

11. Albers, C.A., Lunter, G., MacArthur, D.G., McVean, G., Ouwehand, W.H., and Durbin, R. Dindel: accurate indel calls from short-read data. Genome Res 21, 961-973.

12. Wang, K., Li, M., and Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38, e164.

13. Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29, 308-311.

14. A map of human genome variation from population-scale sequencing. Nature 467, 1061-1073.